Introducing Unobserved Heterogeneity in Earnings Mobility

Thibault Brodaty

INSTITUT NATIONAL DE LA STATISTIQUE ET DES ETUDES ECONOMIQUES
Série des Documents de Travail du CREST
(Centre de Recherche en Economie et Statistique)

n° 2007-35

Introducing Unobserved Heterogeneity in Earnings Mobility

T. BRODATY[1]

[1] Université Paris-1 Panthéon Sorbonne and CREST-INSEE, Paris, France.

Introducing Unobserved Heterogeneity in Earnings Mobility

Thibault BRODATY*

Université Paris-1 Panthéon-Sorbonne & CREST-INSEE

Abstract

This paper introduces and describes unobserved heterogeneity in earnings quintiles transition matrices in the US. Unobserved heterogeneity is found to play a crucial role in earnings mobility. Each individual is attracted, given his characteristics, towards a specific zone of the distribution. At the stationnary equilibrium, the earnings quintiles distribution is thus segmented. Interestingly, while the level of earnings mobility has remained quite stable since 1970, the width of these zones has decreased, such that this segmentation was more pronounced in the 80's and the 90's than in the 70's, especially in the middle of the quintiles distribution.

Résumé

Ce papier introduit de l'hétérogénéité inobservée dans les matrices de transition entre quintiles de salaires. Sa description permet d'étudier la structure de la mobilité des revenus aux Etats-Unis. L'hétérogénéité inobservée joue un rôle essentiel dans la mobilité des revenus et sa caractéristique est que chaque individu est attiré, en fonction de ses caractéristiques, vers une zone spécifique de la distribution des quintiles. A l'équilibre stationnaire, cette distribution est donc segmentée. De plus, alors que le niveau de la mobilité des revenus est resté à peu près stable depuis 1970, la largeur de ces zones a diminué, de sorte que cette segmentation est plus prononcée depuis les années 80, surtout dans le milieu de la distribution.

Keywords: earnings mobility, unobserved heterogeneity, segmentation, state dependence, dynamic multinomial logit.

JEL codes: J31, C33, C35.

---

* thibault.brodaty@ensae.fr; CREST - LMI, 15 Boulevard Gabriel Péri, 92245 Malakoff Cedex, France. Tel: +33(0)1 41 17 59 02. I thank Jean-Marc Robin for the repeated discussions that have significantly improved the quality of this paper. I also thank Bruno Crépon, Thierry Kamionka, Francis Kramarz, Edwin Leuven, Abla Safir, seminar participants at CREST and Paris 1 and conference participants at the 2007 COST meeting on *Evaluation of European Labour Market Programmes* in Sankt Gallen, Journées de Microéconomie Appliquée 2007 in Fribourg, IZA workshop 2007 in Bonn on *Unobserved Heterogeneity in Microeconometric Models*, 2007 RTN meeting on *Micro Data Methods and Practice* in Paris, 2007 AFSE congress in Paris and 2007 European Association of Labour Economists meeting in Oslo for useful suggestions and comments. The usual disclaimer applies.

1 Introduction

Earnings mobility measures the way people move on the income ladder. There are traditionnally two reasons to study such an outcome. First, it is of interest to know in a given economy if it is more or less difficult to move up the economic ladder. In the US, this refers to the very popular idea of the "American dream". Second it is now well known that earnings mobility decreases long term income inequality, which means that a stronger mobility can offset the effects of an increase in cross section inequality[1]. In other words, an increase in earnings inequality (the rich become richer with respect to the poor) is less a concern if mobility increases (that is if the transition poor/rich becomes easier). The case of the United States has been widely studied[2] and the summary of this literature (Gottschalk, 1997) is that there is a substantial degree of mobility, but that when people move they don't move far and that this degree of mobility has not increased in the 80's , which means that earnings mobility has not offset the effects of the dramatic increase of earnings inequality in the 80's.

However, there is a third reason that justifies being interested in the study of earnings mobility, namely the consequences of the structure of earnings mobility on the segmentation of the population. To make this idea clear, let's take a very simple example. Suppose that we are interested in the way workers move, year on year, with respect to the median of the earnings distribution. Suppose that the unconditionnal transition matrix is such that individuals initially below the median have a probability of 0.6 to stay below it (and hence 0.4 to move up) and that individuals initially above the median have a probability of 0.6 to stay above (and hence 0.4 to move down). Suppose now that there is some heterogeneity in the mobility process such that two types of workers exist, each one (50% of the population) being associated with a specific transition matrix, the weighted sum of the conditionnal transition matrices equalling the unconditionnal one. Suppose that type 1 workers have a probability of 1 to stay below and 0.8 to

[1] See Atkinson et al. (1992) for an introduction to earnings mobility.

[2] See for example Buchinsky and Hunt (1999), Gottschalk and Moffit (1998), Burkhauser et al. (1997) and Kopczuk et al. (2007).

move down when they are initially above: when they are below, they don't move and when they are above, most of them move down. In some sense type 1 workers are strongly associated with the bottom of the distribution, they are attracted towards it and it is clear that if the process runs indefinitely, all type 1 workers will be below the median. Now suppose, symmetrically, that type 2 workers have a probability of 1 to stay above and 0.8 to move up when they are initially below: they are attracted towards the top of the distribution and are all above at the stationnary equilibrium. In this very stylised example, it is clear that the introduction of heterogeneity in the mobility process make the population totally segmented at the stationnary equilibrium: type one workers are below the median and type 2 workers are above.

The goal of this paper is to follow this idea: is there some heterogeneity in earnings mobility? And if the answer is positive, what are its consequences on the segmentation of the population? As far as I know this paper is the first to raise this question in these terms, and hence there is no specific related literature. However, some papers on earnings mobility can provide some preliminary answers. For the United States, the issue of obseved heterogeneity has been tackled by Gittleman and Joyce (1995), Schiller (1994) and more recently by Kopczuk et al. (2007). Gittleman and Joyce found that the young, the less educated and blacks are more mobile than those who are older, more educated or white. Schiller (1994) showed that in the 80's young women had more downward mobility and less upward mobility than young men. Kopczuk et al. (2007) confirmed the result that women experience less upward mobility than men. Thus, we know today that the structure of earnings mobility is not the same across demographic groups. In Weber (2002), the author studies whether earnings mobility exhibits state dependence, and shows that if unobserved heterogeneity is not taken into account state dependence is over estimated. This paper is the first to show the presence of unobserved heterogeneity in the mobility process. However, it is not described and hence the question of segmentation can not be addressed.

In this paper, to introduce heterogeneity in quintiles transition matricies, earnings quintiles dynamics is, as in Weber (2002), directly modelled with a dynamic multinomial logit with unobserved heterogeneity. On the one hand, state dependence parameters (that is coefficients associated to lagged dependent variables), being common to all individuals, can be interpreted as labor market flexibility. On the other hand, unobserved heterogeneity, which is individual specific, reflects time-constant, observable (like sex, race and education) and unobservable individual characteristics. The model is estimated in two stages. In the first stage lag coefficients are estimated by conditionnal maximum likelihood (Magnac, 2000), a method that allows the identification to be robust to any specification of the unobserved heterogeneity. In a second stage, and following Heckman and Singer (1984), the support of the unobserved heterogeneity is supposed to be discrete, that is the population is supposed to be composed of types. The initial conditions problem (Heckman, 1981b) is solved with the method proposed by Wooldridge (2005) and the likelihood is maximized via a standard EM algorithm.

The four main results are the following. First, labor market flexibility appears to take the form of a downward rigidity. Second, the population is found to be composed of 5 types in the 70's and 6 in the 80's and the 90's, which means that unobserved heterogeneity plays a major role in earnings mobility: this paper shows that the mobility process is individual specific and that it is not sufficient to compute unconditionnal transition matrices to study earnings mobility. Third, the description of the conditionnal transition matrices shows that the presence of unobserved heterogeneity leads to a segmentation of the quintiles distribution: each individual is strongly associated with a specific zone of the distribution and is attracted towards it, in such a way that at the stationnary equilibrium individuals spend most of their time in that specific zone. The last result is that, interestingly, while the level of earnings mobility has remained quite stable for three decades, its structure has evolved in the middle of the distribution: the specific zones in which individuals move were larger in the 70's, which means that the segmentation was more pronounced in the 80's and the 90's than in the 70's.

The paper is organized as follows. The data is described in section 2, the econometric framework is detailed in section 3, the results are presented in section 4 and section 5 concludes.

2 Data description

2.1 The dataset

I use the Panel Study of Income Dynamics (PSID). The PSID is a representative panel of the US population which was started in 1968 and is still carried out today. The unit of obsevation is the household, the interviews were annual until 1997 and have been bisannual since then. There were 4800 households in 1968 and 7000 in 2001, totaling 65000 individuals observed in 2001. In 1968, the PSID was composed of two separate samples, the first one, the SRC sample, is representative of the US population, whereas the second, the SEO sample, specifically deals with low income households. Thus, low incomes are over-represented in the whole sample and, therefore, I restrict the study to the SRC sample, without using weights, similarly to Moffit and Gottschalk (2002).

To introduce unobserved heterogeneity, panel data are needed. On the one hand, to limit attrition, the step of mobility must be as short as possible; on the other hand the identification of unobserved heterogeneity requires movements in the quintiles distribution and thus a step as large as possible[3]. A step of two years is a good compromise and is adopted in this study. As a consequence, in this paper, earnings mobility must be understood as a short term mobility. In order to minimize measurement error, I use total annual labor income rather than the wage rate because the number of hours worked is poorly reported in the PSID (see for example Bound and al. (1994) for a validation study of the PSID). For the same reason, quintiles are preferred to deciles or ventiles.

The period is splitted into the 70's (1970, 1972, 1974, 1976 and 1978), the 80's (1980, 1982, 1984, 1986 and 1988) and the 90's (1990, 1992, 1994, 1996 and 1998). For each decade I retain individuals who have non-missing observations at each date of the decade. The econometric framework involves data intensive semi-parametric techniques (conditionnal maximum likelihood). To

[3] It is well known in the earnings mobility literature that earnings mobility is an increasing function of its step. It simply reflects that when mobility is measured with a larger step, workers have more time to move and earnings mobility increases.

maximize sample size, I only restrict the sample to 25-60 year-olds at each date of the decade in order to avoid student workers. Moreover this method doesn't allow time-varying variables, that's why the quintiles are not computed directly on the earnings but with the residuals of the regression of the earnings on a constant, age (classified by categories) and age crossed with sex, race and education (College or Not). In order to take into account non-employment, I follow Buchinsky and Hunt (1999) and I create a quintile zero that includes individuals out of the labor market, unemployed or working part-time (at most 1200 hours per year in the 70's, 1300 in the 80's and 1400 in the 90's[4]). Lastly, in order to eliminate outliers, I delete from the decade individuals who achieve at least one time in the decade the transition 1-4, 1-5, 2-5, 5-2, 5-1 or 4-1[5]. The final sample is composed of $1921 * 5$ individual-year in the 70's, $2581 * 5$ in the 80's and $3029 * 5$ in the 90's[6]. Women represent each decade around 55% of the sample (55.1 in the 70's, 54.2 in the 80's and 54.9 in the 90's), Non-white people around 10% (10.0, 9.0 and 8.7), and the share of college graduates is increasing (38.2, 53.0 and 59.1), reflecting the general increase in the education level during the period.

2.2 Descriptive statistics

To describe relative earnings mobility one of the most appropriate tool is the transition matrix between earnings quintiles. Table 1 presents these transition matrices for the 70's, the 80's and 90's. The results are very classical. First, for all periods the most probable transition is the stability. There is more mobility in the middle of the distribution than at the extremes; the vast majority of movements are achieved in an adjacent quintile; upward mobility is slightly greater than downward mobility at the bottom of the distribution and downward mobility is greater than upward mobility at the top. Lastly, earnings mobility has remained fairly stable between

[4] These thresholds have been chosen such that 25% of the workers work less than this value. As the number of hours worked increased between 1970 and 2000, these thresholds have also increased.

[5]The frequence of these transitions is not high enough to be studied within the data intensive econometric framework of this paper.

[6] Sample sizes increase because the PSID follows the children of the original sample individuals

Table 1: Two-Year Quintile Transition matrices, 1970-1998

Origin Quintile	Destination Quintile 0	1	2	3	4	5	Sum	Direction Down	Stable	Up	Sum
70's											
0	84.3	6.8	3.9	2.0	1.6	1.4	100.0	0.0	84.3	15.7	100.0
1	16.3	56.0	20.5	7.2	0.0	0.0	100.0	16.3	56.0	27.7	100.0
2	8.4	17.9	43.4	23.6	6.7	0.0	100.0	26.3	43.4	30.3	100.0
3	6.8	5.3	19.4	40.4	22.3	5.8	100.0	31.5	40.4	28.1	100.0
4	5.1	0.0	5.4	22.5	44.6	22.4	100.0	33.0	44.6	22.4	100.0
5	5.6	0.0	0.0	4.9	21.9	67.6	100.0	32.4	67.6	0.0	100.0
80's											
0	75.0	11.4	6.1	3.7	2.5	1.3	100.0	0.0	75.0	25.0	100.0
1	19.5	53.7	20.6	6.2	0.0	0.0	100.0	19.5	53.7	26.8	100.0
2	11.3	15.4	43.7	22.4	7.2	0.0	100.0	26.7	43.7	29.6	100.0
3	8.9	4.3	16.3	45.2	20.7	4.6	100.0	29.5	45.2	25.3	100.0
4	7.9	0.0	5.4	17.4	46.8	22.5	100.0	30.7	46.8	22.5	100.0
5	4.3	0.0	0.0	5.0	20.3	70.4	100.0	29.6	70.4	0.0	100.0
90's											
0	73.7	11.7	5.6	3.8	2.9	2.3	100.0	0.0	73.7	26.3	100.0
1	21.7	54.9	18.6	4.8	0.0	0.0	100.0	21.7	54.9	23.4	100.0
2	10.5	16.0	47.4	21.2	4.9	0.0	100.0	26.5	47.4	26.1	100.0
3	9.6	3.7	17.7	45.3	19.1	4.6	100.0	31.0	45.3	23.7	100.0
4	8.5	0.0	4.3	17.2	52.6	17.4	100.0	30.0	52.6	17.4	100.0
5	8.3	0.0	0.0	4.4	16.1	71.2	100.0	28.8	71.2	0.0	100.0

Notes: PSID, 25-60 years old, annual earnings. $1921 * 5$ observations in the 70's, $2581 * 5$ in the 80's and $3029 * 5$ in the 90's. Quintile 0 takes into account individuals out of the labor market, unemployed or who work part-time (at most 1200 hours per year in the 70's, 1300 in the 80's and 1400 in the 90's).

the 70's and the 80's and has slightly decreased in the 90's, which implies that earnings mobility has not offset the dramatic increase of earnings inequality on the period.

3 Econometric framework

3.1 Theoritical background

Various classical labor market theories are useful to understand earnings mobility. These models are surveyed in Shiller (1977), Agell and Bennmarker (2007) and Cardoso (2006) and are here briefly summarized.

Earnings stability can be explained by several factors. The first factors are individual specific. The population in each quintile at every period is selected on the basis of their time constant characterisitics, and have thus some propensity to stay in the same quintile in the next period. These characteristics can be linked to the initial investments in general human capital or to some characteristics, like sex and race, that can be sources of discrimination. Time varying characteristics can also generate stability: if those who have the best diplomas accumulate more human capital via experience, then the initial rankings are reinforced. The second factors are common to every individual. Efficiency wages models show that imperfect information on the labor market can lead to downward wage rigidity: firms cannot decrease wages, even if they are hit by a negative demand shock, because workers could decrease their effort or quit. Second institutionnal factors like collective bargaining, employment protection or the existence of a national minimum wage can be the source of earnings immobility.

On the other hand, two factors can explain earnings mobility. The first is productivity shocks, the second is linked to on the job human capital investments: if on the job training implies a sacrifice of present earnings, one might expect movements in relative earnings ranks as individuals experience the burdens and the payoffs of their varying investment decisions.

3.2 The model

To introduce unobserved heterogeneity in earnings mobility, I model quintiles dynamics with a dynamic multinomial logit with unobserved heterogeneity[7]. More precisely, note y_{it} ($y_{it} = 0...5$) the quintile of individual i ($i = 1...N$) at date t ($t = 1...T$). I suppose that

$$\forall k = 0...5\,,\, \forall t = 2...T\,,\; y_{it} = k \text{ if and only if } y^*_{ikt} = \max_{j=0...5}\{y^*_{ijt}\} \text{ where}$$

$$\forall k = 0...5\,,\, \forall t = 2...T\,,\; y^*_{ikt} = \sum_{j=0}^{5} \delta_{jk} \mathbb{1}_{\{y_{it-1}=j\}} + \alpha_{ik} + \epsilon_{ikt} \tag{1}$$

As in every discrete choice model, a reference state must be defined and some parameters set to zero. I choose the non full time employment (quintile 0) and I suppose that $\forall j = 0...5$, $\delta_{j0} = 0$, $\forall i = 1...N$, $\alpha_{i0} = 0$ and $\forall k = 0...5$ $\delta_{0k} = 0$. Because I delete from the sample all the individuals who achieve an extreme transition (1-4, 1-5, 2-5, 5-2, 5-1 or 4-1), these transitions must be impossible in the model and that is why I suppose that $\delta_{14} = \delta_{15} = \delta_{25} = \delta_{41} = \delta_{51} = \delta_{52} = -\infty$. I suppose that the $\epsilon's$ are type I extreme value distributed, and that they are independent accross alternatives, individuals and time, and independent of the α's. I note $\forall i = 1...N$, $\alpha_i = (\alpha_{i0}, ..., \alpha_{i5})'$. Consequently the probability of an individual i being in state k in period t ($t = 2...T$) conditionnal on being in state j in period $t-1$ is given by

$$P(y_{it} = k\,/y_{it-1} = j\,,\,\alpha_i) = \frac{\exp(\delta_{jk} + \alpha_{ik})}{\sum_{l=0}^{5} \exp(\delta_{jl} + \alpha_{il})} \tag{2}$$

which implies that the transition matrix is heterogenous between individuals. Moreover, the law of y_{i1} is supposed to be an unspecified function of α_i, which means that an initial conditions problem must be taken into account.

[7] This specification is preferred to the one that consists in modelling the wages (see Meghir and Pistaferri (2004) and Bonhomme and Robin (2007) for recent examples in this area) and then deriving the quintiles because it is more flexible to model directly the quintiles.

From a statistical point of view, this model can be considered as a "non-linear equivalent" of the ARMA used in the earnings dynamics literature (see for example Burkhauser et al. (1997)): the right hand side of (1) is composed of the lag of the dependent variable (the AR component), a time-constant individual-specific effect and a transitory shock[8]. From an economic point of view, the effect of the lag (ie the state dependence[9]), being common to each individual, can be interpreted as the degree of the flexibility of the labor market. Time-constant individual-specific parameters reflect variables like initial education level, sex, race, motivation, etc[10]. Lastly the transitory component reflects iid productivity shocks.

The interpretation of the parameters is facilitated by writing

$$\frac{P(y_{it} = k \,/ y_{it-1} = j \,,\, \alpha_i)}{P(y_{it} = l \,/ y_{it-1} = j \,,\, \alpha_i)} = \exp((\delta_{jk} - \delta_{jl}) + (\alpha_{ik} - \alpha_{il})) \tag{3}$$

With (3) we can see that the higher $(\alpha_{ik} - \alpha_{il})$, the higher the odds of being in state k with respect to state l, conditionnal on any lagged state j. Now note that if $\alpha_{ik} = \alpha_{il}$ then

$$\frac{P(y_{it} = k \,/ y_{it-1} = j \,,\, \alpha_i)}{P(y_{it} = l \,/ y_{it-1} = j \,,\, \alpha_i)} = \exp(\delta_{jk} - \delta_{jl}) \tag{4}$$

which means that $\delta_{jk} - \delta_{jl}$ is a direct measure of the log odds ratios for observations who have the same "individual propensities" to be in states k and l. It is thus a measure of state dependence and it provides a way to test its presence.

[8] However, papers like Burkhauser et al. find in earnings level dynamics a MA(1) component as well and this AR(1) asumption is thus somewhat restrictive. The study of this potential MA component in earnings quintile dynamics is left for future research.

[9] See Heckman (1981a) for an introduction to state dependence.

[10] This interpretation explains why the individual parameters are state-specific and not the same for all the quintiles: if the effect of education is to increase the probability to be in the top quintile and to decreases the probability to be in the quintile 1, then it cannot enter in the same way in every state equation.

3.3 Estimation

To estimate the model, a two steps procedure is implemented. In the first step the state dependence parameters (the δ's) are estimated with a semi-parametric technique and these estimates are used in a second step to estimate the unobserved heterogeneity (the α's).

In the first step, the conditionnal maximum likelihood method developped by Magnac (2000) is applied. The idea is to condition the likelihood by sufficient statistics such that the conditionnal likelihood does not depend anymore on the individual effects. This method is very attractive because it solves directly the initial conditions problem (Heckman, 1981b) and it allows to make no asumption on the law of unobserved heterogeneity[11]. However with this technique it is not possible to take into account time-varying covariates, which is potentially a problem for time trend and experience. [12]. To solve this problem, the effects of time and experience are eliminated from the earnings: the quintiles are not computed directly on the earnings but with the residuals of the regression of the log-earnings on a constant, age (by class) and age crossed with sex, race and education (College or Not)[13].

Setting Y_{ik} the number of times individual i has been in quintile k between periods 2 and $T-1$, Magnac (2000) shows in his Appendix B that

$$P\left(y_{i2},...,y_{iT-1}\,/y_{i1},Y_{i0},...,Y_{i5},y_{iT}\right)=\frac{\exp\sum_{k>0}\sum_{j}\left(\sum_{t>1}\mathbb{1}_{\{y_{it-1}=j\}}\mathbb{1}_{\{y_{it}=k\}}\delta_{jk}\right)}{\sum_{B_i}\exp\sum_{k>0}\sum_{j}\left(\sum_{t>1}\mathbb{1}_{\{y_{it-1}=j\}}\mathbb{1}_{\{y_{it}=k\}}\delta_{jk}\right)} \tag{5}$$

[11] These features justify the use of a two steps procedure rather than a one step, which would require, to estimate the δ's, to specify the law of the unobserved heterogeneity and to approximate the correlation between the latter and the initial conditions.

[12] To do so the extended method of Honoré and Kyriazidou (2000) could have been used, but the sample size of the PSID is too small for that and it explains why the model has no time varying covariates.

[13] These regressions are performed year by year.

where $B_i = \{b = (y_{i2}, ..., y_{iT-1}) \, / \forall \, k > 0, \, \sum_{t=2}^{T-1} \mathbb{1}_{\{y_{it-1}=j\}} = Y_{ik}\}$ is the set of all possible histories that are compatible with the number of visits to each state from 2 to $T-1$. This conditionnal likelihood doesn't depend anymore on α_i, which means intuitively that individuals who have the same number of visits to each state, but not at the same time, have the same level of unobserved heterogeneity. This is essentially the asumption used to identify the δ's.

To estimate the law of unobserved heterogeneity, ie the law of α_i, asumptions must be done and thus random effects techniques must be applied. To minimize the impact of distributional assumptions I follow Heckman and Singer (1984) and I choose a discrete law, the number of support points (or number of types) being determined with an increasing iterative process stopped when the model has a good fit. To solve the initial conditions problem I adopt the solution proposed by Wooldridge (2005): I maximise the likelihood conditionnal on the initial conditions and I let the law of α_i depend on it. More precisely, I maximize

$$P\left(y_{iT}, ..., y_{i2} \, / y_{i1}, \hat{\delta}\right) = \sum_{l=1}^{L} \left\{ P(\alpha_i = \alpha_i^l / y_{i1}) \prod_{t=2}^{T} P(y_{it}/y_{it-1}, \alpha_i = \alpha_i^l, y_{i1}, \hat{\delta}) \right\} \tag{6}$$

where l denotes one of the L types and $\hat{\delta}$ represents first stage estimates. It must be noted that this likelihood is conditionnal on y_{i1} and that the term $P(\alpha_i = \alpha_i^l / y_{i1})$ shows that the law of α_i depends on y_{i1}. The parameters to be determined are $\forall \, l = 1...L$ and $\forall \, k = 0...5$, α_i^l and $P(\alpha_i = \alpha_i^l / y_{i1} = k)$ (that I note π_{lk}). This likelihood is maximized via a standard EM algorithm, by iterating the two following steps.

- E-Step

For initial values from the step r-1 $(\alpha_i^l)^{(r-1)}$ and $\pi_{lk}^{(r-1)}$, for each type $l = 1...L$ and each individual

i in the sample compute the posterior probability w_{il} that individual i belongs to type l:

$$P(\alpha_i = (\alpha_i^l)^{(r-1)}/y_{iT}, ..., y_{i2}, y_{i1}) = \frac{\pi_{ly_{i1}}^{(r-1)} \prod_{t=2}^{T} P(y_{it}/y_{it-1}, \alpha_i = (\alpha_i^l)^{(r-1)}, y_{i1})}{\sum_{j=1}^{L} \left\{ \pi_{jy_{i1}}^{(r-1)} \prod_{t=2}^{T} P(y_{it}/y_{it-1}, \alpha_i = (\alpha_i^j)^{(r-1)}, y_{i1}) \right\}}$$

- M-step

Update the law of unobserved heterogeneity by averaging the posterior probabilities obtained in the E-step: $\forall\ l = 1...L$, $k = 0...5$,

$$\pi_{lk}^{(r)} = \frac{\sum_{i/y_{i1}=k} w_{il}}{\sum_{i=1}^{N} \mathbb{1}_{\{y_{i1}=k\}}}$$

Then update the support of the unobserved heterogeneity: $\forall\ l = 1...L$,

$$\alpha_{il}^{(r)} = \underset{a}{\operatorname{argmax}} \sum_{i=1}^{N} \sum_{t=2}^{T} w_{il} \ln P(y_{it}/y_{it-1}, \alpha_i = a, y_{i1})$$

The latter expression is the likelihood of a multinomial logit weighted by the E-step posteriors and is thus easily maximized.

The estimation of the model allows me to compute the individual-specific quintiles transition matrix associated with each estimated type of individual using (2). To characterize each transition matrix it is of interest to determine what happens if the process runs indefinitely. To do so I compute, for each matrix, the stationnary equilibrium distributions, that is the quintiles distributions Q such that $Q = M'Q$, where M is the specific transition matrix. Hence, for each transition matrix, the stationnary equilibrium distribution is the eigenvector of the eigenvalue one associated to the transpose of the transition matrix. These distributions show the zone in which individuals move.

4 Results

4.1 Number of types and fit of the model

The number of support points of the unobserved heterogeneity is chosen for every decade in order to maximize the fit to real transition matrices (Table 1). I compare observed transition matrices with simulated ones that I compute using the estimated parameters and (2). The process is iterative in the sens that I start by estimating the model with two types, and if I observe that the fit is not good, I try with three. I stop when the fit is maximum.

For the 70's, the fit with 5 types is better than with 4 and is the same than the one of the model with 6 types. After six types, the algorithm does not converge. To choose between 5 and 6 types (models A and B), I compute the stationnary equilibrium distributions for each type and I observe that both models have 4 types in common while one type of the model A (let's call it type 0) is splitted in two types in the model B (types 00 and 01). In the model A the type 0 is composed of individuals who stay most of the time in the quintile zero and sometimes in the quintile one. In the model B the type 00 is composed of individuals who stay always in the quintile zero, and type 01 of individuals who spend most of their time in quintile zero and sometimes in quintile one. So, types 00 and 01 are a decomposition of the type 0, and therefore models A and B give essentially the same interpretation. As the focus of the paper is primarily the structure of the mobility of individuals who work full time, I don't regard the decomposition of the type associated with the quintile zero as important and I retain the model A. For the 80's, the fits are good both with 5 and 6 types. The problem of the model with 5 types is that one type is composed of individuals who spend their time either in quintile one or in quintile five. This type has no economic interpretation and thus the model with six types is preferred. For the 90's the model with 6 types has the best fit and is thus adopted. The simulated transition matrices are presented in Table 2. They must be compared to the observed ones between brackets. It is clear that the fits of the three models are very good.

Table 2: Fit of the Estimated Models

Origin Quintile	Destination Quintile, 70's						Destination Quintile, 80's						Destination Quintile, 90's					
	0	1	2	3	4	5	0	1	2	3	4	5	0	1	2	3	4	5
0	85.4	6.4	3.7	2.0	1.6	0.9	78.3	9.9	5.2	3.5	1.9	1.1	74.7	11.6	5.1	3.6	2.8	2.2
	(84.3)	(6.8)	(3.9)	(2.0)	(1.6)	(1.4)	(75.0)	(11.4)	(6.1)	(3.7)	(2.5)	(1.3)	(73.7)	(11.7)	(5.6)	(3.8)	(2.9)	(2.3)
1	17.8	55.3	20.1	6.7	0.0	0.0	20.8	54.7	19.0	5.6	0.0	0.0	20.2	56.7	18.7	4.4	0.0	0.0
	(16.3)	(56.0)	(20.5)	(7.2)	(0.0)	(0.0)	(19.5)	(53.7)	(20.6)	(6.2)	(0.0)	(0.0)	(21.7)	(54.9)	(18.6)	(4.8)	(0.0)	(0.0)
2	8.4	18.1	42.4	24.2	7.0	0.0	11.4	16.4	44.8	21.2	6.3	0.0	10.3	16.6	48.1	20.5	4.5	0.0
	(8.4)	(17.9)	(43.4)	(23.6)	(6.7)	(0.0)	(11.3)	(15.4)	(43.7)	(22.4)	(7.2)	(0.0)	(10.5)	(16.0)	(47.4)	(21.2)	(4.9)	(0.0)
3	6.8	5.7	22.1	38.9	20.6	6.0	9.1	4.2	16.6	47.9	18.3	3.8	10.4	3.8	18.1	44.0	19.1	4.6
	(6.8)	(5.3)	(19.4)	(40.4)	(22.3)	(5.8)	(8.9)	(4.3)	(16.3)	(45.2)	(20.7)	(4.6)	(9.6)	(3.7)	(17.7)	(45.3)	(19.1)	(4.6)
4	5.1	0.0	5.7	23.2	43.1	22.9	6.8	0.0	5.7	16.5	49.3	21.7	8.0	0.0	4.8	16.3	52.6	18.3
	(5.1)	(0.0)	(5.4)	(22.5)	(44.6)	(22.4)	(7.9)	(0.0)	(5.4)	(17.4)	(46.8)	(22.5)	(8.5)	(0.0)	(4.3)	(17.2)	(52.6)	(17.4)
5	4.5	0.0	0.0	4.4	22.2	68.9	3.6	0.0	0.0	4.7	20.5	71.2	8.5	0.0	0.0	4.5	16.2	70.8
	(5.6)	(0.0)	(0.0)	(4.9)	(21.9)	(67.6)	(4.3)	(0.0)	(0.0)	(5.0)	(20.3)	(70.4)	(8.3)	(0.0)	(0.0)	(4.4)	(16.1)	(71.2)

Notes: PSID, 25-60 years old, annual earnings. Between brackets observed quintiles transition matrices. The simulated ones are computed using the estimated parameters and (2).

4.2 State dependence

The estimates of state dependence parameters (the δ's) are presented in Table 3. To interpret the results, the relevant measure is the difference between the δ's and not the δ's themselves. For example, if $\delta_{21} < \delta_{22}$, an individual initially in the quintile 2 and who has $\alpha_{i1} = \alpha_{i2}$[14] has a greater probability to stay than to move down in quintile one. Thus the comparison between the δ's can be interpreted as a comparison of probabilities for people who have the same "characteristic-related" propensities. Moreover, these coefficients being common to all individuals, they can be interpreted as a measure of labor market flexibility. The first result of Table 3 is that all coefficients are positive and significantly different from zero: labor market flexibility makes the transition towards quintile zero the less probable one. Now if I concentrate on the quintiles one to five it is interesting to test the equality of the parameters on each row. These (Wald) tests are presented in Table 4. The first remark is that almost all the results are similar for the three decades. The hypothesis that parameters from a same row are equal is rejected for all the rows, which means that there is some state dependence in earnings mobility. This result is consistent with Weber (2002). To see if the state dependence comes from upward or downward mobility I perform tests on some parts of each row. On each row, the hypothesis that the parameter δ_{jj} is equal to the parameters δ_{jk}, $k > j$, can not be rejected. On the contrary, the hypothesis that the parameter δ_{jj} is equal to the parameters δ_{jk}, $k < j$ is strongly rejected. In other words, the state dependence favors upward mobility as much as the stability, but more than downward mobility. In some sense, the state dependence prevents from moving down. Therefore, as far as state dependence is a measure of labor market flexibility, the latter seems to behave like a downward rigidity. This result refers to the literature on wage rigidity, and in particular on downward wage rigidity, in which it is now well known that wages are rigid, and especially downward (see for example Agell and Benmarker for a recent study). This paper shows that it is also true for earnings quintiles.

[14] That is someone who has the same "characteristic-related" propensities to be in the quintiles 1 and 2.

Table 3: State Dependence Parameters

Origin Quintile	Destination Quintile, 70's						Destination Quintile, 80's						Destination Quintile, 90's					
	0	1	2	3	4	5	0	1	2	3	4	5	0	1	2	3	4	5
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
1	0	1.6	1.9	2.0	$-\infty$	$-\infty$	0	1.5	1.3	1.2	$-\infty$	$-\infty$	0	1.6	1.5	1.0	$-\infty$	$-\infty$
	–	(0.24)	(0.26)	(0.35)	–	–	–	(0.17)	(0.19)	(0.25)	–	–	–	(0.16)	(0.18)	(0.24)	–	–
2	0	1.5	2.7	2.9	2.6	$-\infty$	0	0.9	1.8	1.8	1.8	$-\infty$	0	1.3	2.2	2.2	1.7	$-\infty$
	–	(0.29)	(0.32)	(0.33)	(0.39)	–	–	(0.20)	(0.21)	(0.22)	(0.27)	–	–	(0.19)	(0.21)	(0.21)	(0.27)	–
3	0	1.1	2.4	3.4	3.1	3.0	0	0.5	1.5	2.5	2.5	2.5	0	0.4	1.6	2.4	2.5	2.5
	–	(0.35)	(0.33)	(0.39)	(0.39)	(0.48)	–	(0.26)	(0.23)	(0.26)	(0.27)	(0.37)	–	(0.24)	(0.21)	(0.24)	(0.24)	(0.32)
4	0	$-\infty$	2.3	3.5	3.7	4.0	0	$-\infty$	1.4	2.1	3.5	3.9	0	$-\infty$	1.3	2.0	3.0	2.9
	–	–	(0.41)	(0.40)	(0.45)	(0.50)	–	–	(0.27)	(0.27)	(0.33)	(0.39)	–	–	(0.26)	(0.24)	(0.26)	(0.29)
5	0	$-\infty$	$-\infty$	2.7	3.7	4.6	0	$-\infty$	$-\infty$	2.2	3.8	4.8	0	$-\infty$	$-\infty$	1.5	2.1	2.9
	–	–	–	(0.51)	(0.51)	(0.60)	–	–	–	(0.37)	(0.38)	(0.46)	–	–	–	(0.29)	(0.28)	(0.31)

Notes: PSID, 25-60 years old, annual earnings. For each decade, the parameter crossing the origin quintile j and the destination quintile k is the estimate of δ_{jk}. Standard errors between brackets.

Table 4: Wald Tests of the Presence of State Dependence

H0	DF	70's Wald Statistics	70's P-value	80's Wald Statistics	80's P-value	90's Wald Statistics	90's P-value
$\delta_{11} = \delta_{12} = \delta_{13}$	2	1.86	0.39431	1.80	0.40748	6.46	0.03959
$\delta_{11} = \delta_{12}$	1	1.45	0.22895	0.88	0.34882	0.33	0.56336
$\delta_{21} = \delta_{22} = \delta_{23} = \delta_{24}$	3	24.70	0.00002	26.00	0.00001	28.13	0.00000
$\delta_{22} = \delta_{23} = \delta_{24}$	2	1.14	0.56436	0.03	0.98738	3.73	0.15469
$\delta_{22} = \delta_{23}$	1	0.71	0.40028	0.02	0.90179	0.02	0.87511
$\delta_{22} = \delta_{21}$	1	18.73	0.00002	22.11	0.00000	23.45	0.00000
$\delta_{31} = \delta_{32} = \delta_{33} = \delta_{34} = \delta_{35}$	4	41.50	0.00000	56.53	0.00000	71.89	0.00000
$\delta_{33} = \delta_{34} = \delta_{35}$	2	1.61	0.44656	0.04	0.98021	0.01	0.99577
$\delta_{33} = \delta_{34}$	1	1.28	0.25877	0.01	0.91994	0.01	0.92715
$\delta_{31} = \delta_{32} = \delta_{33}$	2	40.09	0.00000	48.18	0.00000	62.26	0.00000
$\delta_{32} = \delta_{33}$	1	11.31	0.00077	16.63	0.00005	15.42	0.00009
$\delta_{42} = \delta_{43} = \delta_{44} = \delta_{45}$	3	18.08	0.00042	61.80	0.00000	42.70	0.00000
$\delta_{44} = \delta_{45}$	1	0.38	0.53636	1.26	0.26079	0.22	0.63914
$\delta_{42} = \delta_{43} = \delta_{44}$	2	16.26	0.00029	51.51	0.00000	40.11	0.00000
$\delta_{43} = \delta_{44}$	1	0.55	0.45945	28.04	0.00000	18.06	0.00002
$\delta_{53} = \delta_{54} = \delta_{55}$	2	15.23	0.00049	41.74	0.00000	18.61	0.00009
$\delta_{54} = \delta_{55}$	1	4.77	0.02902	9.84	0.00171	8.98	0.00274

Notes: PSID, 25-60 years old, annual earnings.

4.3 Unobserved heterogeneity

With the estimates of the δ's and the α's and the expression (2) of the transition probabilities deduced from the model, I compute a specific quintile transition matrix for each type of individuals (ie for each discrete value of α). These matrices are presented for the 70's, the 80's and the 90's in the Tables 6, 7 and 8 in the appendix. Let's look for example at the specific matrix in the bottom left of Table 8. This type represents 12.7% of the population. When these individuals are initially in quintile 3, the most probable destination quintile is the quintile 3 (61.1%). When they are initially in quintile 2, the most probable transition is also to move up in the quintile 3 (53.5%), and in fact for all the origin quintiles (but the quintile zero, which is a particular quintile), the quintile 3 is the most probable destination. Individuals of this type can be characterized by the fact that they are "attracted" towards the quintile 3, whichever quintile they start in. That is why this type can be labelled type 3. Actually this phenomenon is present for all the types of Table 8. For example, in the top right matrix, it is clear that the column corresponding to the destination quintile 2 is the mode of each conditionnal distribution[15], which means that this specific transition matrix attracts towards the quintile 2 and can be labelled type 2. These results mean that in the 90's, each individual, given his time-constant characteristics, is attracted towards a specific quintile. This is also true in the 80's (see Table 7 in the appendix), but the situation is different in the 70's. Let's look for example at the bottom right transition matrix in Table 6. Individuals of this type are clearly "attracted" towards the quintile 5 and this type is thus called type 5. This type is similar to those of the 80's and the 90's. This is also the cases of the so-called types 0 and 1. But if we look for example at the top right matrix, we see that when the individuals come from the bottom of the distribution (quintiles 1, 2 and to some extent 3) they take the direction of quintile 2, whereas when they are initially in the quintiles 4 or 5 they arrive in the quintile 3[16]. It is still true that individuals are "attracted" towards a zone of the

[15] Note that it is not true for the origin quintile 5. It stems from the fact that, by hypothesis, the transition 5-2 is not allowed ($\delta_{52} = -\infty$), and thus the probability of the transition 5-2 equals zero. The most probable destination quintile is then the adjacent quintile 3.

[16] This type is thus called type 2-3

distribution, but that zone does'nt seem to be restricted to a specific quintile. This phenomenon is also true for the so-called type 3-4.

To see more precisely the zones that individuals are attracted to, in other words the zones to which individuals, at the stationnary equilibrium, move most of the time, it is useful to compute, for each type-specific matrix, the quintiles stationnary equilibrium distributions[17]. These distributions are presented in Table 5. For example, for the 90's, at the stationnary equilibrium, we see that 8.4% of the type 2 individuals are in the quintile 0, 10.9%, 55.6%, 22.4%, 2.7% and 0.0% in the quintiles 1, 2, 3, 4 and 5 respectively. This means that a majority of type 2 individuals are each year, at the stationnary equilibrium, in the quintile 2, or in other words that type 2 individuals spend most of their time, at the stationnary equilibrium, in quintile 2. Table 5 shows that this result is true for all types, for the 80's and the 90's. This confirms the idea that the nature of the heterogeneity is to attract individuals towards a specific quintile in the 80's and the 90's. Thus the presence of heterogeneity in the earnings mobility process leads to a segmentation of the population. The meaning of that segmentation, in the 80's and the 90's, is that each individual spends most of the time in the same quintile, revolving around from time to time. For the 70's, type 2-3 individuals spend at the stationnary equilibrium respectively 42.1% and 35.1% of their time in the quintiles 2 and 3, and type 3-4 individuals 29.4% and 49.3% in the quintiles 3 and 4. These two types of individuals spend most of their time in two quintiles. The difference between the 70's and the following decades is that, in the middle of the distribution in the 70's, the zones are not centered around a specific quintile but are rather composed of two quintiles that are both regularly visited. Thus in the middle of the distribution, the zones in which individuals move are larger and have a greater intersection in the 70's, which means that the segmentation was less pronounced at that time.

As a robustness check I change the definition of the quintile zero, by including only zero wages (out of the labor market or unemployed all the year), part-time workers being taken into

[17] See section 3.3.

Table 5: Specific Quintiles Stationnary Distributions

70's

Type	Quintile 0	1	2	3	4	5	Sum
0	83.7	5.0	4.1	3.7	2.5	1.0	100.0
1	14.5	61.9	18.8	2.4	0.0	2.4	100.0
$2-3$	6.3	8.2	42.1	35.1	6.4	1.9	100.0
$3-4$	5.1	3.9	4.6	29.4	49.3	7.7	100.0
5	10.4	0.0	0.3	4.1	20.2	65.0	100.0

80's

Type	Quintile 0	1	2	3	4	5	Sum
0	89.1	6.9	2.1	0.9	0.8	0.3	100.0
1	22.2	63.7	11.5	2.2	0.4	0.0	100.0
2	2.9	24.3	61.9	9.3	1.5	0.2	100.0
3	11.9	4.3	18.7	48.4	16.1	0.6	100.0
4	8.9	2.5	3.1	14.0	48.7	22.8	100.0
5	14.7	5.4	15.2	4.1	6.3	54.4	100.0

90's

Type	Quintile 0	1	2	3	4	5	Sum
0	90.5	5.8	2.0	1.2	0.2	0.3	100.0
1	27.4	52.3	15.3	4.0	1.0	0.0	100.0
2	8.4	10.9	55.6	22.4	2.7	0.0	100.0
3	11.9	3.6	12.8	52.8	17.6	1.3	100.0
4	15.4	1.0	3.5	13.3	51.1	15.7	100.0
5	11.1	1.0	2.1	2.0	8.9	75.0	100.0

Notes: PSID, 25-60 years old, annual earnings.

account to compute the other quintiles. But due to small sample size in the 70's and the 80's (as stated earlier, conditionnal maximum likelihood is data intensive), the δ's can not be identified, and thus this check is performed only for the 90's. The specific quintiles distributions at the stationnary equilibrium are presented in Table 9 in the appendix. The results are qualitatively the same: there is a clear segmentation of the population: each type of individual spends most of the time in a specific quintile.

5 Conclusion

This paper studies the structure of earnings mobility in the United States. The idea is to decompose a quintiles transition matrix into a weighted sum of individual-specific matrices and to describe the nature of these specific matrices. To achieve this decomposition, unobserved heterogeneity is introduced in earnings quintiles transition matrices. I model quintiles dynamics directly as a dynamic multinomial logit with unobserved heterogeneity, which allows me to disentangle labor market flexibility from time-constant individual characteristics. In a first stage, in order to compute estimates robust to any specification of the unobserved heterogeneity, state dependence parameters are estimated with a semi-parametric conditionnal maximum likelihood. In a second stage the support of the unobserved heterogeneity is supposed to be discrete, the initial conditions problem is taken into account and the likelihood is maximized via a standard EM algorithm.

In the first stage, the estimation of the state dependence parameters shows that, since it decreases downward mobility probabilities, labor market flexibility seems to be a downward rigidity. Then, the estimation of the entire model shows, by identifying 5 or 6 types of individual-specific transition matrices, that unobserved heterogeneity plays a crucial role in earnings mobility. The nature of these matrices is to "attract" each individual towards a zone of the quintiles distribution, and thus the effect of this heterogeneity is to segment the population. Interestingly, while the level of earnings mobility has remained quite stable since 1970, its structure has changed in the middle of the quintiles distribution: the specific zones in which individuals move were larger in the 70's than in the 80's and the 90's. This result shows that via the presence of unobserved heterogeneity, earnings mobility segments the population and that this segmentation was more pronounced in the 80's and the 90's than in the 70's.

references

Agell, J. and H. Benmarker (2007). Wage Incentives and Wage Rigidity: A Representative View from Within, *Labour Economics*, 14, 347-369.

Atkinson, A., Bourguignon, F. and Morisson, C. (1992). Empirical Studies of Earnings Mobility, *Philadelphia: Harwood Academic Publisher.*

Bonhomme, S. and Robin, J.M. (2007). Assessing the Equalizing Force of Mobility Using Short panels: France, 1990-2000, *mimeo*

Bound, J., Brown, C., Duncan, G. and Rodgers, W. (1994). Evidence on the validity of cross-sectional and longitudinal labor market data, *Journal of Labor Economics*, 12, 3, 345-368.

Buchinsky, M. and Hunt, J. (1999). Wage mobility in the United-States, *The Review of Economics and Statistics*, 81, 3, 351-368.

Burkhauser, R., Hotz, D. and Rhody, S. (1997). Labour earnings mobility and inequality in the United-States and Germany during the growth years of the 1980s, *International Economic Review*, 38, 4, 775-794.

Cardoso, A.R. (2006). Wage Mobility: Do Institutions Make a Difference?, *Labour Economics*, 13, 387-404.

Gittleman, M., and M. Joyce (1995). Earnings Mobility in the United States, 1967-1991, *Monthly Labor Review*, 118, 9, 3-13.

Gottschalk, P. (1997). Inequality, income growth, and mobility: the basic facts, *The Journal of Economic Perspectives*, 11, 2, 21-40.

Heckman, J.J. (1981a). Heterogeneity and state dependence, *in Studies in Labor Markets, ed. by*

S. Rosen, pp. 91-139. Chicago University Press, Chicago.

Heckman, J.J. (1981b). The Incedental Paramter Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process, *in Structural Analysis of Discrete Data with Econometric Applications, ed. by C. Manski, and D.McFadden*, pp. 179-195. MIT Press, Cambridge, MA.

Heckman, J.J., and B. Singer (1984). A Method For Minimizing the Impact of Distributional Assumptions in Econometric Models For Duration Data, *Econometrica*, 52, 271-320.

Honore, B.E. and Kyriazidou, E. (2000). Panel Data Discrete Choice Models with Lagged Dependent Variables, *Econometrica*, 68, 4, 839-874.

Kopczuk, W., Saez, E. and J. Song (2007). Uncovering the American Drean: Inequality and Mobility in Social Security Earnings Data Since 1937, *NBER Working Paper* No 13345.

Magnac, T. (2000). Subsidized training and youth employment, *The Economic Journal*, 110, 805-37.

Meghir, R. and Pistaferri, L. (2004) Income Variance Dynamics and Heterogeneity, *Econometrica*, 72, 1-32.

Moffit, R. and Gottshalk, P. (1998). Trends in the variance of permanent and transitory earnings in the United-States and their relation to earnings mobility, *mimeo.*

Moffit, R. and Gottshalk, P. (2002). Trends in the transitory variance of earnings in the United-States, *The Economic Journal*, 112, C68-C73.

Schiller, B. (1977). Relative Earnings Mobility in the United States, *American Economic Review*, 67, 5, 926-941.

Schiller, B. (1994). Relative Earnings Redux: Youth Mobility in the 80's, *Review of Income and Wealth*, 40, 441-456.

Weber, A. (2002) . State Dependence and Wage Dynamics: A Heterogeneous Markov Chain Model for Wage Mobility in Austria, *Economics Series*, 114, Institute for Advanced Studies.

Wooldridge, J. (2005). Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity, *Journal of Applied Econometrics*, 20, 1, 39-54.

Table 6: Type-Specific Transition Matrices, 70's

Origin	Dest Q, Type 0 (38.6%)							Dest Q, Type 1 (12.9%)							Dest Q, Type 2-3 (19.3%)						
Quintile	0	1	2	3	4	5	Sum	0	1	2	3	4	5	Sum	0	1	2	3	4	5	Sum
0	90.2	4.3	2.4	1.5	1.2	0.4	100.0	38.0	50.1	9.1	0.9	0.0	1.9	100.0	35.2	16.4	30.1	13.7	3.3	1.2	100.0
1	65.0	15.3	11.9	7.9	0.0	0.0	100.0	10.7	69.9	17.4	1.9	0.0	0.0	100.0	8.3	19.3	48.4	24.0	0.0	0.0	100.0
2	48.1	10.3	18.8	14.7	8.2	0.0	100.0	9.2	54.8	31.9	4.1	0.0	0.0	100.0	4.2	8.8	51.7	30.2	5.1	0.0	100.0
3	42.8	6.2	13.1	22.1	11.7	4.1	100.0	10.5	42.3	28.8	8.0	0.0	10.3	100.0	3.7	5.2	35.9	45.3	7.4	2.6	100.0
4	38.2	0.0	10.0	20.5	21.1	10.3	100.0	14.6	0.0	34.1	11.5	0.0	39.8	100.0	3.6	0.0	29.7	45.5	14.4	6.9	100.0
5	44.9	0.0	0.0	10.8	23.2	21.2	100.0	16.3	0.0	0.0	5.7	0.0	77.9	100.0	7.2	0.0	0.0	41.1	27.2	24.4	100.0

Origin	Destination Quintile							Dest Q, Type 3-4 (13.8%)							Dest Q, Type 5 (15.3%)						
Quintile	0	1	2	3	4	5	Sum	0	1	2	3	4	5	Sum	0	1	2	3	4	5	Sum
0								38.4	16.7	5.6	13.7	22.8	2.9	100.0	67.9	0.1	0.8	3.6	10.3	17.3	100.0
1								14.8	31.8	14.5	38.9	0.0	0.0	100.0	67.7	0.4	5.6	26.3	0.0	0.0	100.0
2								5.1	10.1	10.8	34.0	39.9	0.0	100.0	24.1	0.1	4.2	23.6	47.9	0.0	100.0
3								3.4	4.5	5.6	38.4	43.0	5.0	100.0	9.1	0.0	1.2	15.1	29.1	45.4	100.0
4								2.3	0.0	3.2	26.8	58.2	9.4	100.0	4.3	0.0	0.5	7.4	27.8	59.9	100.0
5								2.7	0.0	0.0	14.1	63.9	19.3	100.0	3.1	0.0	0.0	2.4	18.8	75.7	100.0

Notes: PSID, 25-60 years old, annual earnings.

Table 7: Type-Specific Transition Matrices, 80's

Origin	Dest Q, Type 0 (26.1%)							Dest Q, Type 1 (11.6%)							Dest Q, Type 2 (8.4%)						
Quintile	0	1	2	3	4	5	Sum	0	1	2	3	4	5	Sum	0	1	2	3	4	5	Sum
0	91.1	5.8	1.6	0.7	0.6	0.2	100.0	43.1	46.9	8.0	1.5	0.5	0.0	100.0	10.5	34.0	47.9	6.3	1.1	0.2	100.0
1	72.8	20.6	4.8	1.8	0.0	0.0	100.0	15.1	72.9	10.3	1.7	0.0	0.0	100.0	2.9	42.0	49.5	5.6	0.0	0.0	100.0
2	74.2	11.4	8.1	3.6	2.7	0.0	100.0	19.9	52.0	22.3	4.4	1.5	0.0	100.0	2.5	19.0	67.9	9.1	1.5	0.0	100.0
3	72.7	7.4	6.1	6.7	5.3	1.8	100.0	24.0	41.6	20.8	10.0	3.6	0.0	100.0	2.8	14.3	59.3	19.4	3.3	0.8	100.0
4	69.2	0.0	4.9	4.6	14.5	6.8	100.0	40.6	0.0	29.5	12.3	17.5	0.0	100.0	3.5	0.0	62.6	17.8	12.1	4.0	100.0
5	63.5	0.0	0.0	4.7	16.6	15.1	100.0	53.5	0.0	0.0	17.8	28.6	0.1	100.0	7.3	0.0	0.0	40.8	31.5	20.4	100.0
Origin	Dest Q, Type 3 (20.8%)							Dest Q, Type 4 (18.1%)							Dest Q, Type 5 (15.0%)						
Quintile	0	1	2	3	4	5	Sum	0	1	2	3	4	5	Sum	0	1	2	3	4	5	Sum
0	38.8	9.3	18.3	26.9	6.4	0.2	100.0	44.8	11.3	6.2	13.8	18.7	5.3	100.0	50.1	12.3	21.5	3.5	2.4	10.2	100.0
1	16.6	17.7	29.0	36.7	0.0	0.0	100.0	27.7	31.0	14.2	27.1	0.0	0.0	100.0	25.6	28.0	40.7	5.7	0.0	0.0	100.0
2	10.2	5.9	29.4	44.4	10.1	0.0	100.0	14.6	8.8	12.2	27.9	36.5	0.0	100.0	20.3	12.0	52.9	8.9	5.9	0.0	100.0
3	7.2	2.8	16.0	59.1	14.4	0.5	100.0	8.2	3.3	5.3	29.6	41.3	12.2	100.0	13.5	5.3	27.2	11.1	7.9	34.9	100.0
4	6.8	0.0	12.5	40.1	38.6	1.9	100.0	4.1	0.0	2.2	10.7	59.1	23.9	100.0	6.6	0.0	11.2	4.0	11.1	67.2	100.0
5	6.5	0.0	0.0	42.6	46.4	4.5	100.0	2.8	0.0	0.0	8.0	49.8	39.5	100.0	3.5	0.0	0.0	2.3	7.3	86.9	100.0

Notes: PSID, 25-60 years old, annual earnings.

Table 8: Type-Specific Transition Matrices, 90's

Origin	Dest Q, Type 0 (19.6%)							Dest Q, Type 1 (20.5%)							Dest Q, Type 2 (15.1%)						
Quintile	0	1	2	3	4	5	Sum	0	1	2	3	4	5	Sum	0	1	2	3	4	5	Sum
0	92.3	4.8	1.5	0.9	0.2	0.3	100.0	50.6	36.8	8.9	2.6	1.0	0.0	100.0	30.2	15.7	39.1	13.1	1.9	0.0	100.0
1	73.8	18.9	5.3	1.9	0.0	0.0	100.0	18.2	65.1	14.2	2.5	0.0	0.0	100.0	9.6	24.6	54.9	11.0	0.0	0.0	100.0
2	69.2	13.5	10.4	6.1	0.8	0.0	100.0	16.9	45.9	27.5	7.9	1.9	0.0	100.0	5.2	10.2	62.5	20.3	1.8	0.0	100.0
3	75.0	5.8	6.1	8.6	1.9	2.6	100.0	26.3	28.5	23.2	15.9	6.1	0.0	100.0	7.2	5.6	46.2	35.9	5.2	0.0	100.0
4	81.5	0.0	4.8	6.0	3.4	4.3	100.0	41.3	0.0	26.5	16.0	16.2	0.0	100.0	9.9	0.0	46.3	31.7	12.1	0.0	100.0
5	89.7	0.0	0.0	4.1	1.5	4.7	100.0	71.7	0.0	0.0	17.2	11.1	0.0	100.0	28.8	0.0	0.0	57.2	13.9	0.1	100.0

Origin	Dest Q, Type 3 (12.7%)							Dest Q, Type 4 (18.7%)							Dest Q, Type 5 (13.4%)						
Quintile	0	1	2	3	4	5	Sum	0	1	2	3	4	5	Sum	0	1	2	3	4	5	Sum
0	40.2	8.3	12.5	30.1	8.3	0.7	100.0	50.5	3.6	5.3	10.7	23.3	6.7	100.0	44.5	4.8	6.6	2.4	7.2	34.6	100.0
1	18.7	18.9	25.7	36.8	0.0	0.0	100.0	42.2	14.7	19.5	23.6	0.0	0.0	100.0	42.9	22.6	28.3	6.2	0.0	0.0	100.0
2	8.0	6.2	23.0	53.5	9.3	0.0	100.0	14.9	4.0	14.4	28.2	38.5	0.0	100.0	24.0	9.6	33.0	11.7	21.6	0.0	100.0
3	7.1	2.2	11.0	61.1	17.2	1.4	100.0	9.0	1.0	4.7	22.2	49.2	13.9	100.0	7.4	1.2	5.5	4.7	14.1	67.2	100.0
4	8.3	0.0	9.3	45.8	34.0	2.5	100.0	6.9	0.0	2.6	10.9	63.5	16.0	100.0	5.3	0.0	2.9	2.1	17.1	72.6	100.0
5	15.7	0.0	0.0	54.0	25.5	4.8	100.0	12.6	0.0	0.0	12.3	45.7	29.4	100.0	6.2	0.0	0.0	1.5	7.8	84.5	100.0

Notes: PSID, 25-60 years old, annual earnings.

Table 9: Specific Quintile Stationnary Distributions in the 90's, Quintile Zero Only Composed of Zero Earnings.

	Quintile						
Type	0	1	2	3	4	5	Sum
0	72.7	16.6	5.4	2.5	1.2	1.6	100.0
1	12.9	66.0	16.5	3.9	0.7	0.0	100.0
2	4.3	17.3	63.6	12.9	1.4	0.5	100.0
3	5.4	5.9	16.8	55.4	14.3	2.2	100.0
4	4.3	3.5	3.1	11.9	60.9	16.2	100.0
5	4.3	1.4	1.1	1.5	6.5	85.1	100.0

Notes: PSID, 25-60 years old, annual earnings.